Preliminary GPT-5-mini v2 Analysis of Reddit Empathy/Support Annotations

Yi Zhang

June 16, 2026

Contents

Research Questions

This project asks how support-seeking interactions unfold in online Reddit conversations. The analyses are organized around three linked questions:

1. How do people seek support online? That is, what support-seeking needs are commonly expressed in support-seeking posts?
2. How do other users respond? In particular, are comment strategies **tuned** to the type of support being sought in the original post?
3. Do higher-fit responses (e.g., providing emotional support for emotional disclosure posts) lead to more positive OP uptake, such as gratitude, continued elaboration, follow-up questions, or potentially other engagement signals such as upvotes?

Dataset

The current analysis is based on a Reddit corpus of support/advice-seeking posts and comments. The full dataset contains over 3000 posts from 24 subreddits, with over 300K comments. For the present project, I created a stratified subset of 2765 post-comment-reply pairs from about 1000 posts across the 24 subreddits.

Subreddit Distribution

Table 1: Subreddit distribution in completed annotations

subreddit	n	percent
teenagers	264	9.5%
datingoverthirty	249	9.0%
relationships	244	8.8%
Teachers	187	6.8%
Parenting	170	6.1%
socialskills	163	5.9%
NoStupidQuestions	154	5.6%
DecidingToBeBetter	136	4.9%
AskWomenOver30	135	4.9%
askwomenadvice	115	4.2%
Advice	107	3.9%

subreddit	n	percent
AskGaybrosOver30	106	3.8%
AskParents	103	3.7%
AskMenOver30	96	3.5%
needadvice	85	3.1%
AskMenAdvice	81	2.9%
findapath	79	2.9%
teaching	57	2.1%
badroommates	52	1.9%
love	46	1.7%
AskReddit	44	1.6%
AskHR	39	1.4%
AskOldPeopleAdvice	36	1.3%
thingsmykidsaid	17	0.6%

Annotation and Measurement Approach

Codebook development

Each post–comment–OP reply unit was annotated with three families of multi-label binary codes. **Support-seeking need** labels describe what kind of response the original post appears to invite, including advice/problem-solving, emotional disclosure, validation or appraisal, sense-making, high distress, high-stakes urgency, and perspective-sharing/not clearly support-seeking. **Comment strategy** labels describe what the level-1 reply does, including emotional acknowledgment, validation/affirmation, reflection/paraphrase, interpretation, reframing, challenge/criticism, question/exploration, advice/problem-solving, and self-disclosure. **OP uptake** labels describe how the original poster responds when they reply, including gratitude, elaboration, answering a question, follow-up questions, and pushback.

The codebook was developed and refined iteratively based on LLM annotation and human audits (based on a human-annotated gold standard of ~100 posts). The current codebook and prompt (updated 6/14/2026) was able to achieve good reliability (high precision and recall) for a majority of the labels, including advice-seeking, validation/appraisal, sense-making, interpretation, question/exploration, self-disclosure, gratitude, and elaboration. Some labels remain less reliable, especially high distress, high stakes, reflection/paraphrase, reframing, challenge, and low-frequency OP uptake categories. These labels should therefore be treated as exploratory and interpreted with caution.

The label set is grounded in established social support and supportive-message traditions that distinguish emotional support, informational support, appraisal/validation, validation/normalization, and instrumental support/problem-solving functions (e.g., Cutrona & Suhr, 1992; Barbee & Cunningham, 1995; Goldsmith, 2004). The response-move labels also align with recent computational work on empathic language tactics, such as validation, paraphrasing, reappraisal, and self disclosure (Gueorguieva et al., 2026). Finally, we also annotated OP uptake of the comments because recent work argues that effective support on Reddit cannot be determined by the comments alone but should also take into account how the comments are received by the OP and the community (Alghamdi et al., 2025). The current project extends this work by explicitly modeling the fit between what the OP appears to seek, what the commenter provides, and how the OP responds.

Codebook Summary

Support-Seeking Need Labels

Table 2: Support-seeking need labels from the current codebook

Label	Definition	Examples/cues
Advice/problem-solving invited	OP invites advice, next steps, solutions, coping strategies, or help deciding what to do.	What should I do? Any advice? How should I handle this?
Emotional disclosure/venting	OP's emotional experience is clear, central, and directly related to why they are posting.	Fear, shame, grief, loneliness, anger, sadness, relief, vulnerability.
Validation/reassurance/appraisal invited	OP seeks appraisal, reassurance, normalization, or judgment about their own feeling, reaction, decision, behavior, concern, or situation.	Am I wrong? Am I overreacting? Is this normal? Was this reasonable?
Sense-making/confusion	OP invites help understanding meaning, causes, motives, conflicting feelings, relationship dynamics, or how to interpret a specific situation.	Why did this happen? What does this mean? I feel torn.
High distress/calming need	OP expresses acute severe distress or a clear need for calming or grounding.	Panic, spiraling, hopelessness, inability to function.
High-stakes/urgent concern	The post involves concrete serious risk or urgent real-world consequences.	Self-harm, abuse, medical/legal/housing/job crisis, minors at risk, immediate safety.
Perspective-sharing/not support-seeking	Boundary flag for posts mainly sharing an update, reflection, PSA, gratitude, personal experience, or general discussion prompt rather than seeking support for OP's own problem.	Update post, lesson learned, general community opinion-seeking.

Comment Strategy Labels

Table 3: Comment strategy labels from the current codebook

Label	Definition	Examples/cues
Emotional acknowledgment/sympathy	Comment directly expresses care, concern, sympathy, condolence, or emotional recognition toward OP or OP's situation.	I'm sorry; that sounds hard; my condolences.
Validation/affirmation	Comment validates, reassures, normalizes, affirms, encourages, or empowers OP's feelings, perspective, concerns, choices, efforts, worth, or ability.	You're not overreacting; that makes sense; you're allowed to set a boundary.
Reflection/paraphrase/understanding	Comment explicitly restates, mirrors, or summarizes OP's situation, feelings, dilemma, or perspective.	You're dealing with X; the fact that Y happened sounds tough.
Interpretation/explanation	Comment gives the commenter's take on what is happening, why it is happening, motives/responsibilities, or the underlying problem.	It sounds like he is avoiding responsibility; this may be happening because...

Label	Definition	Examples/cues
Reframing/alternative perspective	Comment actively shifts how OP might appraise, define, or emotionally understand the situation.	This does not mean you failed; focus on what you can control.
Challenge/confrontation/criticism	Comment challenges OP's own framing, assumptions, expectations, or behavior.	You're being unfair; I think you're missing something.
Question/exploration	Comment asks a genuine information-seeking, clarification-seeking, reflection-inviting, or conversation-continuing question.	Have you talked to them? What feels hardest?
Advice/problem-solving	Comment offers concrete or general recommendations, next steps, coping strategies, or behavioral guidance.	Talk to them; set a boundary; consider therapy.
Self-disclosure/personal experience	Commenter shares their own lived experience, identity, background, similar situation, personal history, or relevant first-person experience.	When I went through this; as someone who...

OP Uptake Labels

Table 4: OP uptake labels from the current codebook

Label	Definition	Examples/cues
Gratitude/acknowledgment	OP thanks, appreciates, or explicitly acknowledges the comment as useful, meaningful, or received.	Thank you; this helps; good point.
Elaboration/continued disclosure	OP adds meaningful context, reasoning, emotion, reflection, or continuation of the story.	Additional background, feelings, or reflection beyond a short factual answer.
Answering the commenter's question	OP mainly answers a question asked by the commenter.	Direct answer to a genuine question.
Follow-up question/asks for more	OP asks for clarification, more advice, more explanation, or asks a new question.	What do you think I should say?
Pushback/correction/disagreement	OP clearly corrects, resists, disagrees with, or rejects the commenter's point.	That's not what I meant; I don't think that's true.

Summary of key findings

This report summarizes a preliminary analysis of 2765 completed GPT-5-mini v2 annotations from a planned 3,000-case sample of Reddit post–comment–OP reply units.

The sample is intentionally enriched for OP replies, gratitude/positive uptake proxies, high engagement, question-like comments, and long/information-rich cases. The `sample_random == 1` subset is retained as a benchmark random core. Therefore, full-sample base rates are useful for diagnostics and exploratory modeling, but should not be interpreted as population prevalence.

The strongest preliminary pattern is that support-seeking needs and comment response moves align in theoretically sensible ways. Advice-seeking posts are much more likely to receive advice/problem-solving

comments; emotional disclosure posts are more likely to receive emotional acknowledgment; and sense-making posts are more likely to receive interpretation/explanation. Comment questions strongly predict OP replies, elaboration, and answering-question uptake; challenge predicts OP pushback.

Sample and Annotation Diagnostics

Table 5: Sample and annotation diagnostics

metric	value
planned_sample_rows	3000
completed_annotations	2765
parse_missing_or_unfinished	235
error_rows	26
unique_completed_comments	2765
unique_completed_posts	1180
unique_subreddits	24
with_op_reply_completed	964

Human Audit Reliability

The table below summarizes the results of the most recent round of human audit, in which the experimenter manually labeled 50 cases as the gold standard. Because this validation sample is small, these numbers are best interpreted as a check for obvious systematic problems rather than final reliability estimates.

The strongest current labels include advice-seeking, validation/appraisal, interpretation, question/exploration, self-disclosure, gratitude, and elaboration. Labels with lower support or weaker agreement, especially high stakes, high distress, reflection/paraphrase, reframing, and some low-frequency OP uptake categories, should be treated more cautiously.

Sampling Strata

Because the current analyses are based on a stratified (rather than completely randomly selected) sample, I keep track of how each post-comment-reply pair is selected. That is, whether a particular comment was selected because it was random chosen from the full corpus, or because it was OP-replied, gratitude-like, high-engagement, question-like, or long/information-rich. The table below summarizes the proportion of posts in each category.

Table 6: Completed annotations by sample stratum

stratum	n	percent
sample_random	1398	50.6%
sample_op_replied	449	16.2%
sample_gratitude_proxy	231	8.4%
sample_high_engagement	225	8.1%
sample_question_like	228	8.2%
sample_long_info_rich	237	8.6%
sample_fill	31	1.1%

Comparing random and stratified samples

The table below compares the features between comments that are completely randomly sampled vs. stratified.

Table 7: Completed random benchmark subset compared with full enriched sample

sample	n	posts	OP reply	Mean comment words	Median comment words	Mean post words
Full enriched	2765	1180	34.9%	64.8	34	358.7
Random benchmark	1398	792	10.3%	54.4	28	268.8

Label Base Rates

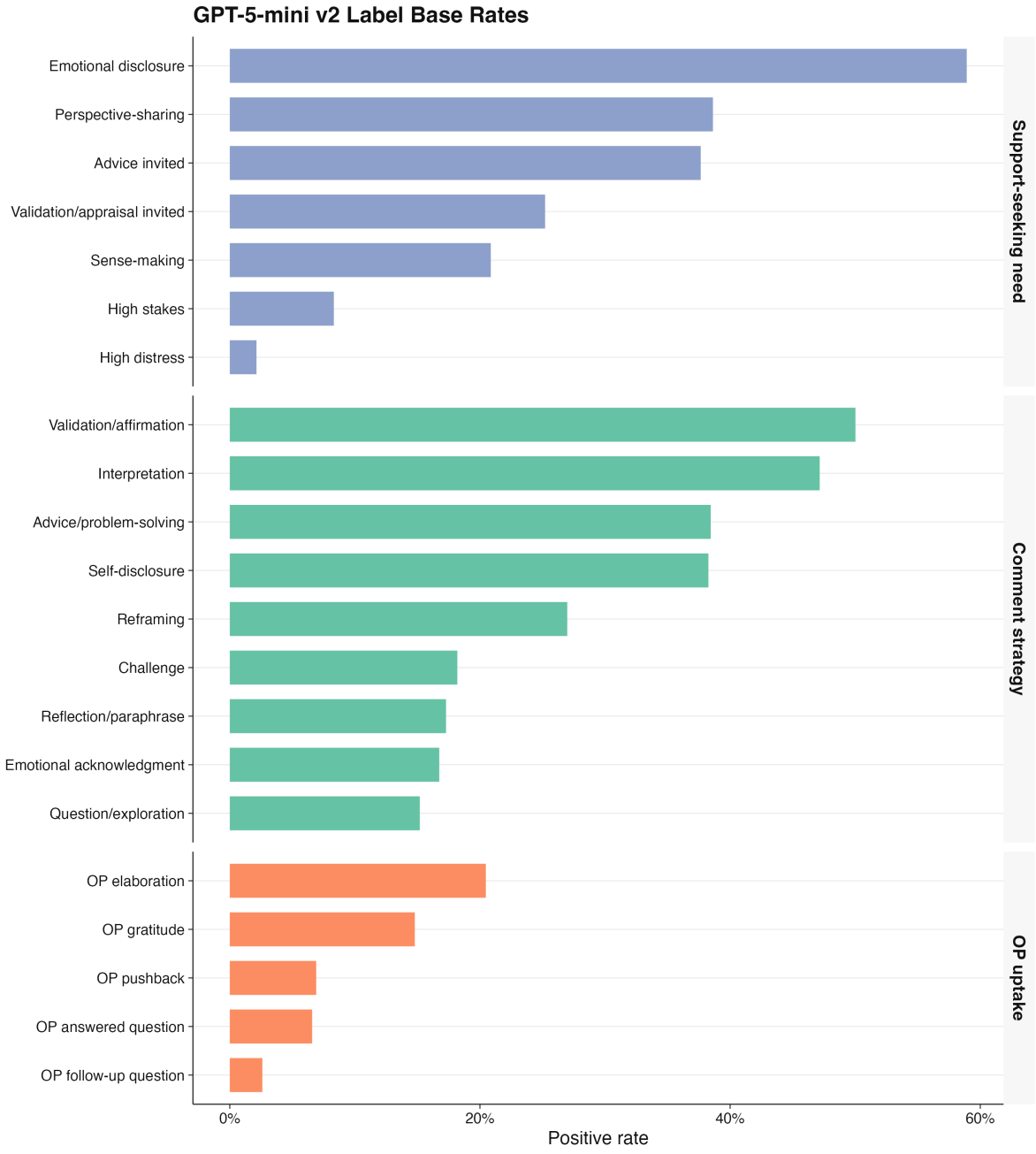


Table 8: LLM-coded label base rates

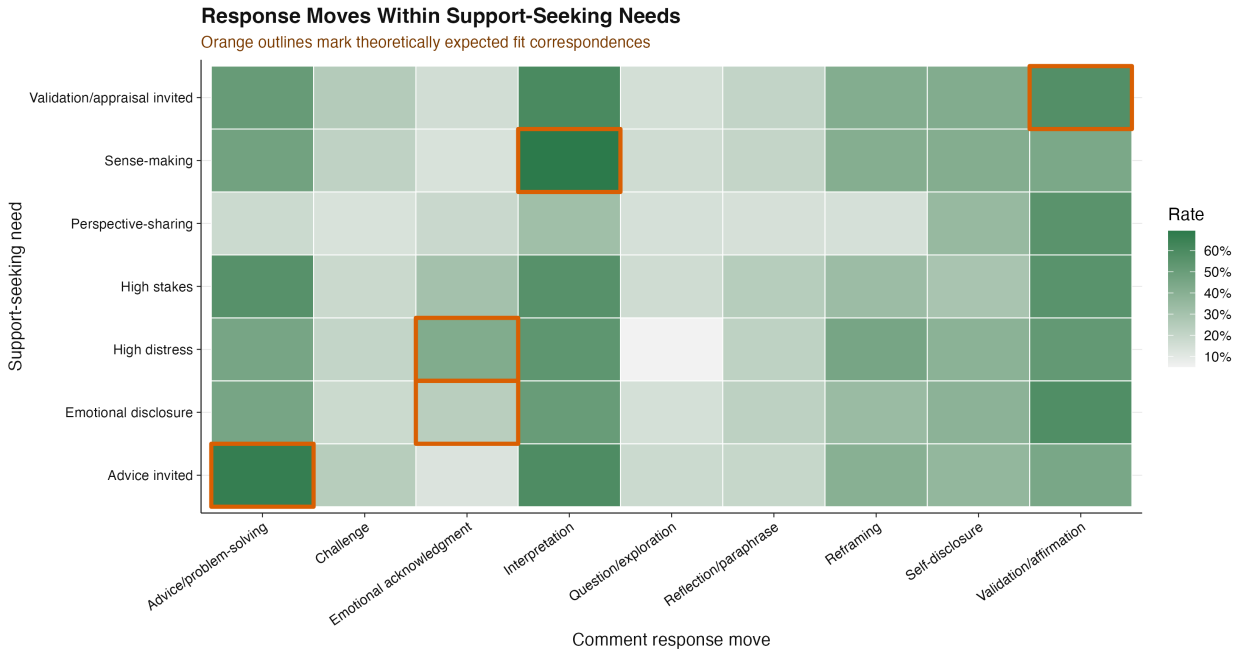
group	label	positives	Full rate	Random rate
Post	Advice invited	1041	37.6%	33.6%
Post	Emotional disclosure	1629	58.9%	54.9%
Post	Validation/appraisal invited	697	25.2%	25.4%
Post	Sense-making	577	20.9%	20.2%
Post	High distress	59	2.1%	2.3%

group	label	positives	Full rate	Random rate
Post	High stakes	230	8.3%	7.5%
Post	Perspective-sharing	1068	38.6%	38.1%
Comment	Emotional acknowledgment	463	16.7%	15.6%
Comment	Validation/affirmation	1383	50.0%	49.9%
Comment	Reflection/paraphrase	478	17.3%	14.7%
Comment	Interpretation	1304	47.2%	45.3%
Comment	Reframing	746	27.0%	24.2%
Comment	Challenge	503	18.2%	15.9%
Comment	Question/exploration	420	15.2%	9.3%
Comment	Advice/problem-solving	1063	38.4%	35.3%
Comment	Self-disclosure	1058	38.3%	38.0%
OP uptake	OP gratitude	409	14.8%	2.9%
OP uptake	OP elaboration	566	20.5%	5.4%
OP uptake	OP answered question	182	6.6%	2.4%
OP uptake	OP follow-up question	72	2.6%	0.9%
OP uptake	OP pushback	191	6.9%	2.7%

The most common support-seeking needs were emotional disclosure, perspective-sharing/not clearly support-seeking, and advice-seeking. The most common comment moves were validation/affirmation, interpretation/explanation, advice/problem-solving, and self-disclosure. OP elaboration and gratitude were the most common uptake labels.

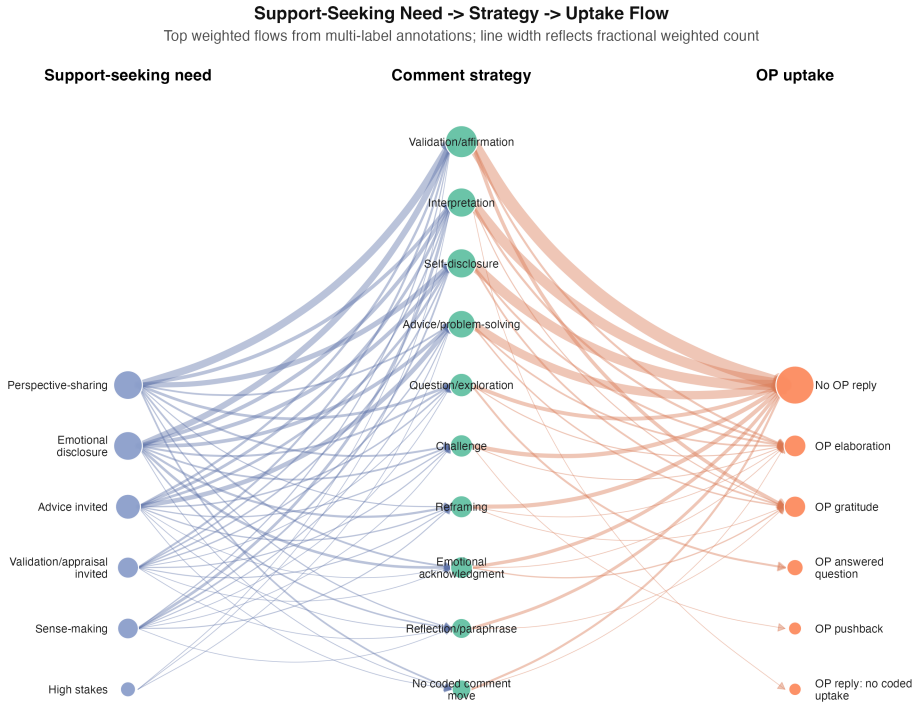
Do user responses track the OP's support-seeking needs?

We first examine whether replies appear responsive to the type of support being sought in the original post. For example, when a post asks for advice, do commenters provide advice/problem-solving? When a post discloses distress or emotion, do commenters provide emotional acknowledgment or validation? These descriptive rates provide the first construct check before moving to mixed-effects models.



The heatmap shows strong construct coherence: advice-seeking posts receive advice/problem-solving responses at high rates, emotional disclosure posts receive more acknowledgment/validation, and sense-making posts receive more interpretation and questions.

Flow chart of support seeking and provision dynamics



This flow chart characterizes the dynamics of support seeking and provision: which OP support-seeking needs lead to which comment responses strategies and which strategies lead to which OP uptakes.

Mixed-Effects Models

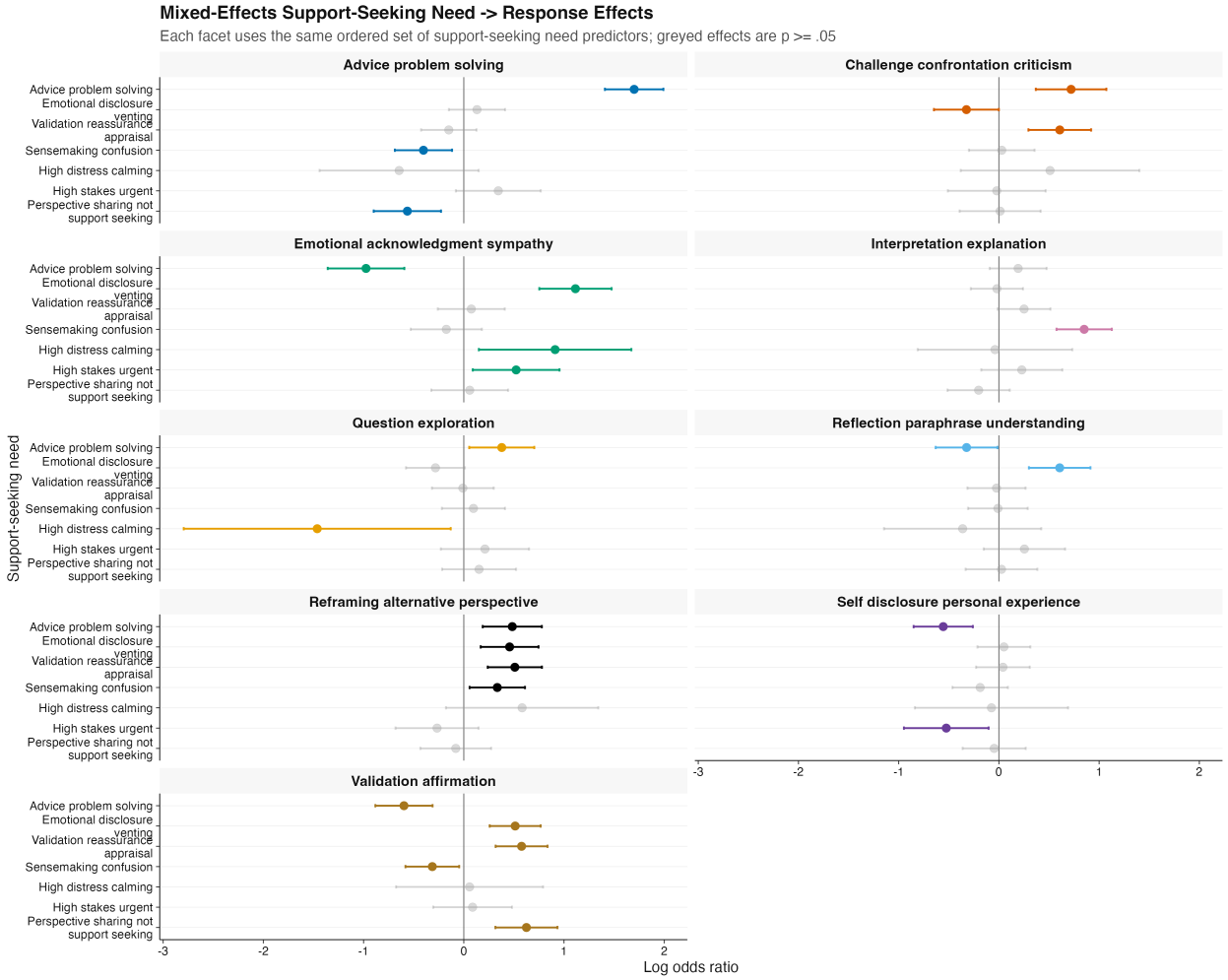
Next, I ran mixed-effects models to formally test whether support-seeking needs predict response strategies, while accounting for comment/post features such as word count and upvotes. These models ask whether commenters choose different strategies depending on the type of support sought by the OP. For example, if a post primarily invites emotional support, are commenters more likely to provide emotional acknowledgment or validation? If a post asks for advice, are commenters more likely to give advice?

Models use logistic mixed-effects models where feasible. The preferred random-effects structure was `(1 | subreddit_id/post_id)`, and was simplified to `(1 | subreddit_id)` to prevent singular fit when needed. These random effects help account for clustering among comments within posts and subreddits.

Table 9: Largest support-seeking need-to-response mixed-effects associations

outcome	predictor	OR [95% CI]	p
Comment advice problem solving	Post advice problem solving	5.47 [4.09, 7.33]	<.001
Comment question exploration	Post high distress calming	0.23 [0.06, 0.88]	0.031

outcome	predictor	OR [95% CI]	p
Comment emotional acknowledgment sympathy	Post emotional disclosure venting	3.05 [2.12, 4.37]	<.001
Comment emotional acknowledgment sympathy	Post advice problem solving	0.38 [0.26, 0.55]	<.001
Comment emotional acknowledgment sympathy	Post high distress calming	2.49 [1.16, 5.32]	0.019
Comment interpretation explanation	Post sensemaking confusion	2.34 [1.78, 3.09]	<.001
Comment challenge confrontation criticism	Post advice problem solving	2.05 [1.45, 2.92]	<.001
Comment advice problem solving	Post high distress calming	0.52 [0.24, 1.16]	0.111
Comment validation affirmation	Post perspective sharing not support seeking	1.87 [1.37, 2.55]	<.001
Comment challenge confrontation criticism	Post validation reassurance appraisal	1.84 [1.34, 2.51]	<.001
Comment reflection paraphrase understanding	Post emotional disclosure venting	1.83 [1.35, 2.49]	<.001
Comment validation affirmation	Post advice problem solving	0.55 [0.41, 0.73]	<.001



Does comment response strategy predict OP uptake?

The next set of analyses asks whether different response strategies are associated with different forms of OP uptake. These models test, for example, whether questions invite OP elaboration or answers, whether validation is associated with gratitude, and whether challenge or criticism is associated with OP pushback.

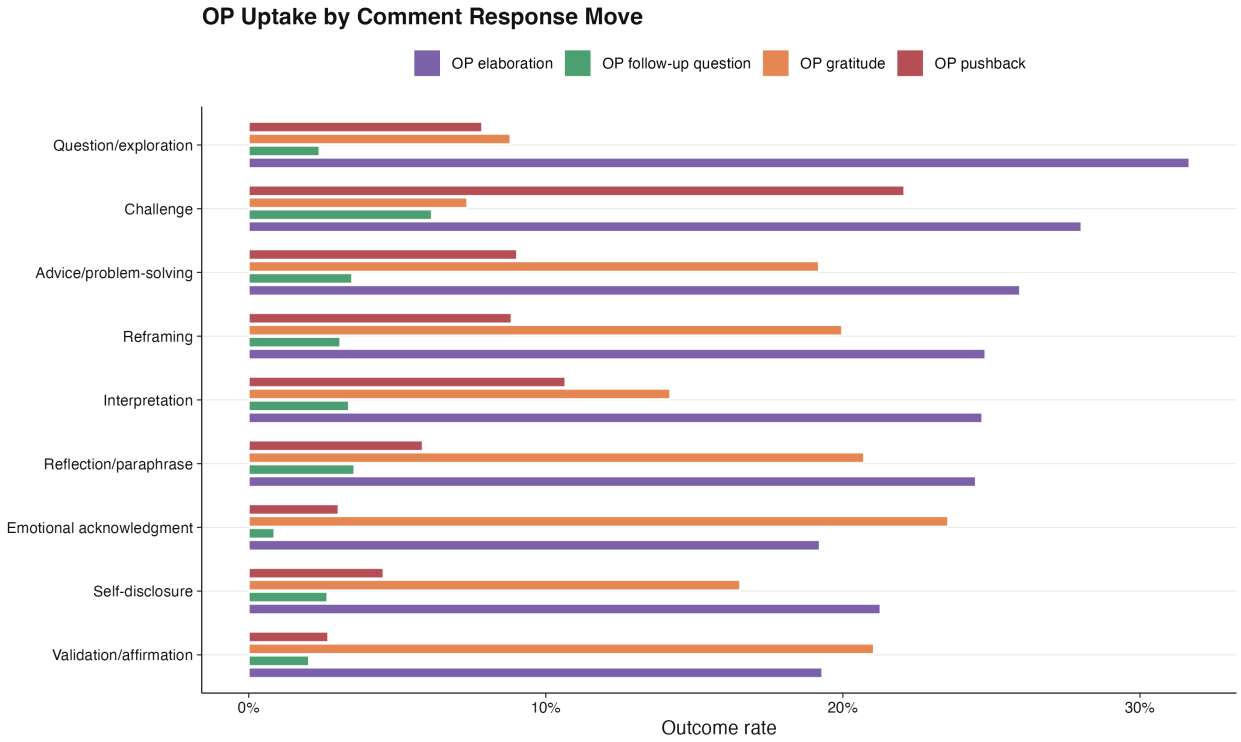
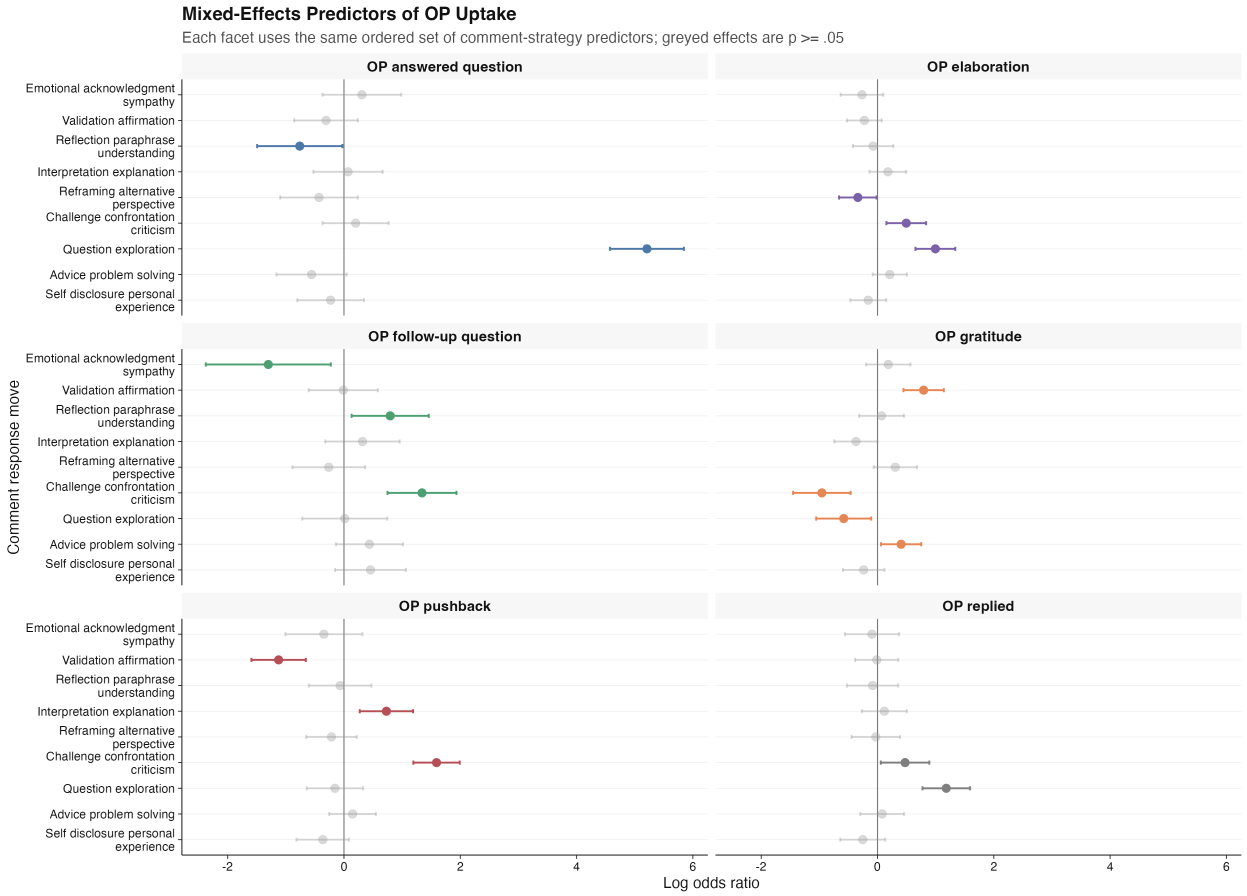


Table 10: Largest response-to-uptake mixed-effects associations

outcome	predictor	OR [95% CI]	p
Op answering question	Comment question exploration	183.26 [97.04, 346.07]	<.001
Op pushback correction disagreement	Comment challenge confrontation criticism	4.91 [3.30, 7.32]	<.001
Op followup question asks for more	Comment challenge confrontation criticism	3.83 [2.11, 6.93]	<.001
Op followup question asks for more	Comment emotional acknowledgment sympathy	0.27 [0.09, 0.80]	0.018
Op replied any	Comment question exploration	3.26 [2.17, 4.91]	<.001
Op pushback correction disagreement	Comment validation affirmation	0.33 [0.20, 0.52]	<.001
Op elaboration continued disclosure	Comment question exploration	2.70 [1.92, 3.81]	<.001
Op gratitude acknowledgment	Comment challenge confrontation criticism	0.38 [0.23, 0.63]	<.001
Op followup question asks for more	Comment reflection paraphrase understanding	2.21 [1.14, 4.30]	0.019
Op gratitude acknowledgment	Comment validation affirmation	2.21 [1.56, 3.13]	<.001

outcome	predictor	OR [95% CI]	p
Op answering question	Comment reflection paraphrase understanding	0.47 [0.22, 0.97]	0.042
Op pushback correction disagreement	Comment interpretation explanation	2.07 [1.31, 3.28]	0.002



Question/exploration is the strongest predictor of OP answering a question and is also associated with OP reply and elaboration. Challenge/confrontation predicts OP pushback. Validation/affirmation is positively associated with gratitude and negatively associated with pushback.

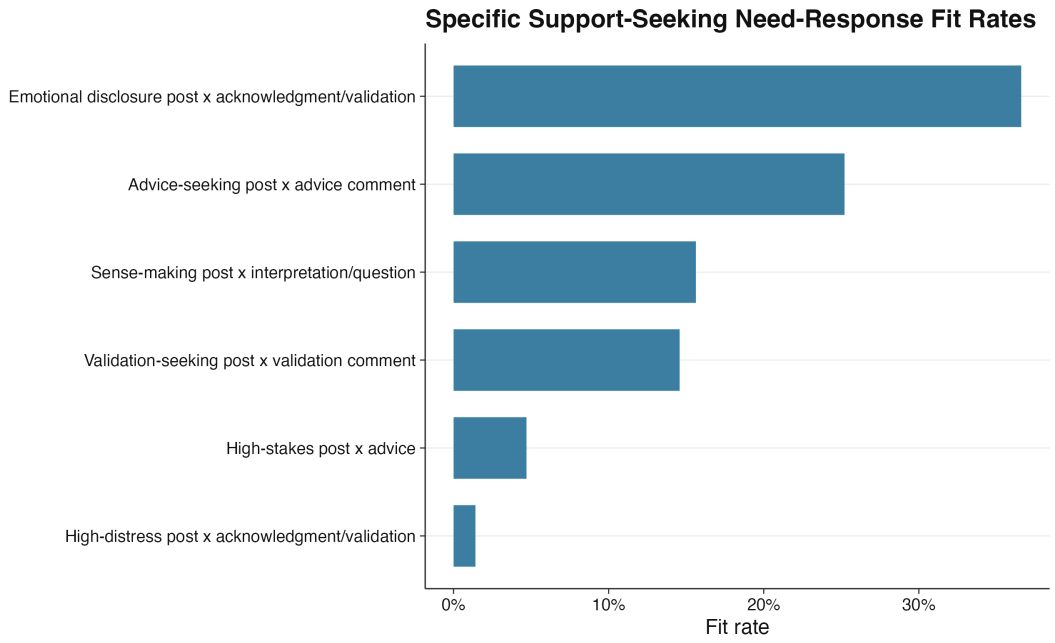
Does the fit between support seeking and provision predict OP uptake?

To understand whether greater fit between support-seeking and provision is associated with more positive OP uptake, I specified, for each support-seeking need, whether it received a good-fit response. Fit is defined based on the following paring:

- Advice seeking post + advice/problem-solving focused response
- Emotional disclosure post + validation
- Emotional disclosure post + emotional acknowledgement
- Sense-making post + interpretation/explanation

- Validation/appraisal seeking post + validation

The key question is whether high-fit responses predict OP gratitude, elaboration, follow-up, or lower push-back.



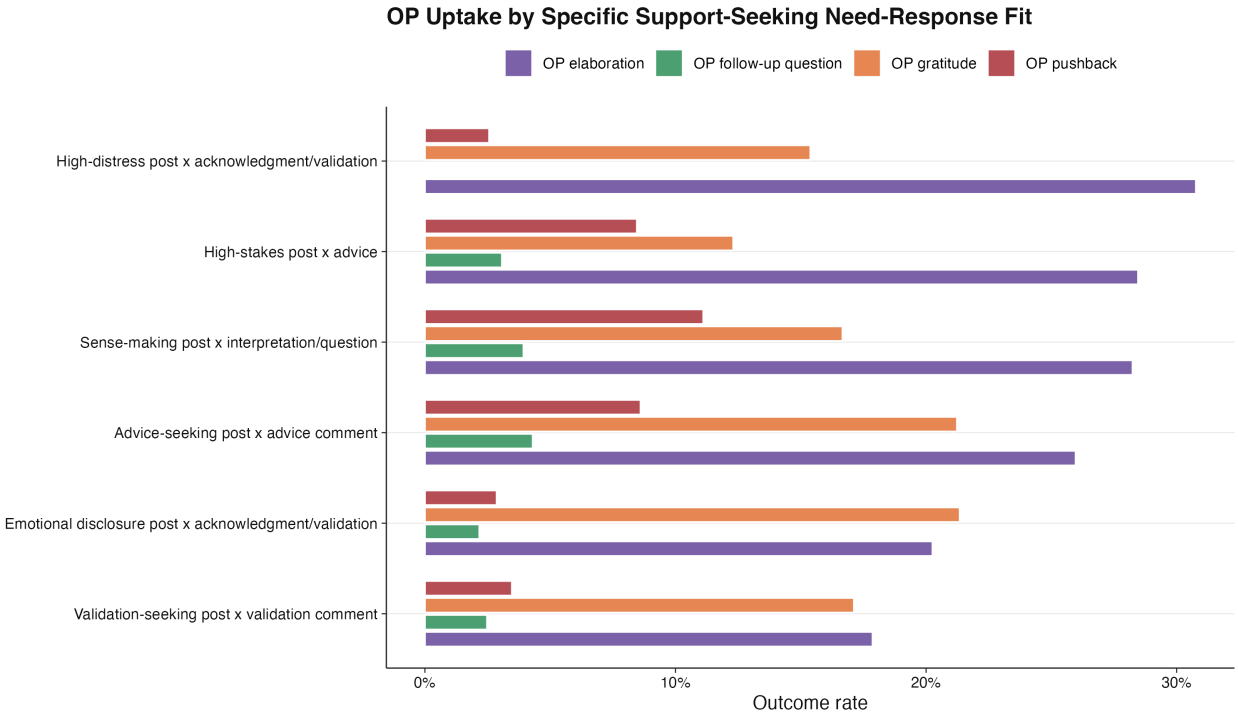


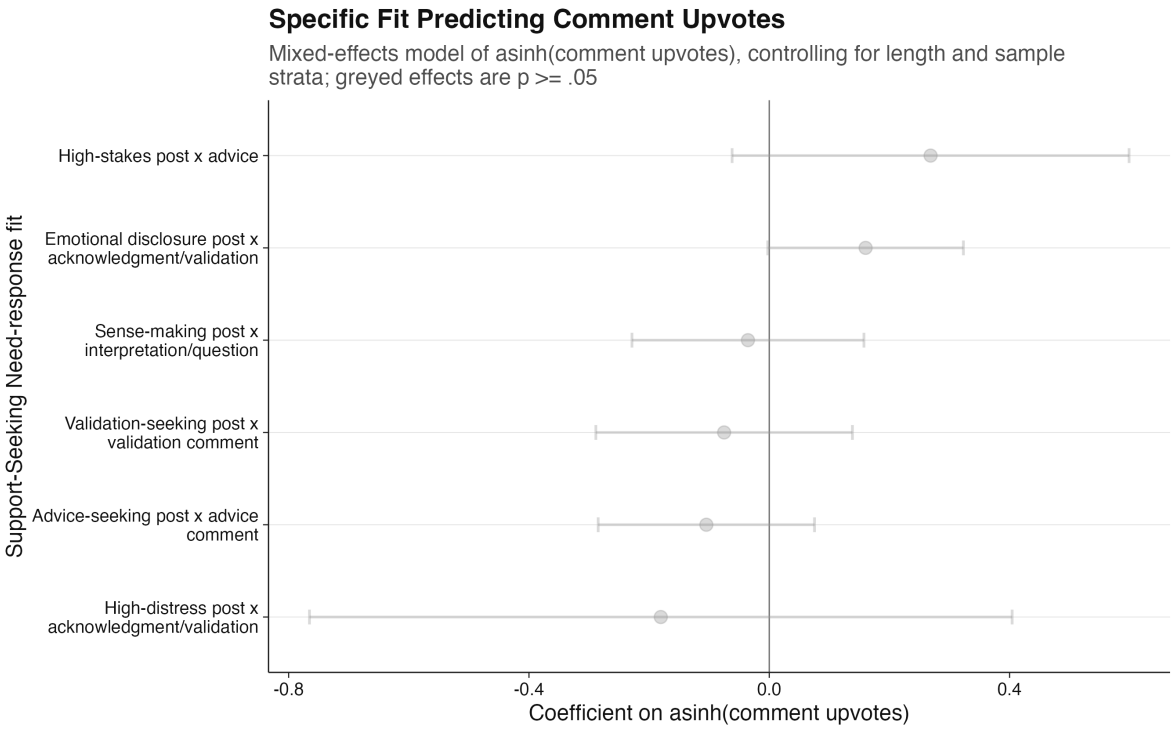
Table 11: Specific fit indicators predicting OP uptake

outcome	predictor	OR [95% CI]	p
Op followup question asks for more	Fit distress ack validation	0.00 [0.00, Inf]	0.992
Op pushback correction disagreement	Fit disclosure ack validation	0.23 [0.14, 0.38]	<.001
Op gratitude acknowledgment	Fit highstakes advice	0.30 [0.13, 0.66]	0.003
Op gratitude acknowledgment	Fit disclosure ack validation	2.27 [1.60, 3.21]	<.001
Op followup question asks for more	Fit advice advice	2.10 [1.20, 3.67]	0.009
Op gratitude acknowledgment	Fit advice advice	2.01 [1.39, 2.92]	<.001
Op pushback correction disagreement	Fit sensemaking interpret question	1.89 [1.24, 2.90]	0.003
Op elaboration continued disclosure	Fit distress ack validation	1.84 [0.64, 5.30]	0.258
Op followup question asks for more	Fit sensemaking interpret question	1.63 [0.89, 2.99]	0.117
Op elaboration continued disclosure	Fit validation validation	0.63 [0.41, 0.97]	0.037
Op elaboration continued disclosure	Fit disclosure ack validation	0.64 [0.47, 0.86]	0.004
Op elaboration continued disclosure	Fit sensemaking interpret question	1.48 [1.05, 2.09]	0.024
Op gratitude acknowledgment	Fit distress ack validation	0.70 [0.20, 2.51]	0.585
Op gratitude acknowledgment	Fit validation validation	0.72 [0.46, 1.15]	0.169
Op followup question asks for more	Fit disclosure ack validation	0.74 [0.40, 1.35]	0.320
Op pushback correction disagreement	Fit validation validation	0.74 [0.38, 1.44]	0.380
Op pushback correction disagreement	Fit advice advice	1.25 [0.83, 1.89]	0.291
Op followup question asks for more	Fit highstakes advice	0.82 [0.27, 2.48]	0.732
Op pushback correction disagreement	Fit distress ack validation	0.85 [0.10, 7.26]	0.881
Op pushback correction disagreement	Fit highstakes advice	1.15 [0.52, 2.54]	0.731
Op elaboration continued disclosure	Fit advice advice	1.12 [0.81, 1.55]	0.504
Op elaboration continued disclosure	Fit highstakes advice	1.09 [0.60, 1.97]	0.777
Op gratitude acknowledgment	Fit sensemaking interpret question	0.93 [0.61, 1.42]	0.746
Op followup question asks for more	Fit validation validation	0.94 [0.43, 2.08]	0.883

Preliminary patterns suggest that advice-seeking/advice fit and disclosure/acknowledgment-validation fit are associated with gratitude, while sense-making fit can also be associated with more OP elaboration and pushback, indicating that interpretive/question responses may invite further clarification or correction.

Does fit between support-seeking and provision predict upvotes?

As an exploratory analysis, we also tested whether specific support-seeking need-response fit predicts the number of upvotes received by the level-1 comment. Because Reddit upvotes are highly skewed, the model predicts `asinh(comment_upvotes)`, which behaves similarly to a log transform for large positive values while preserving zero values. These models include the same specific fit indicators, control for post/comment length and sample strata, and use mixed-effects structure where feasible.


Specific Fit Predicting Comment Upvotes
Mixed-effects model of asinh(comment upvotes), controlling for length and sample strata; greyed effects are p >= .05
Support-Seeking Need-response fit
High-stakes post x advice
Emotional disclosure post x acknowledgment/validation
Sense-making post x interpretation/question
Validation-seeking post x validation comment
Advice-seeking post x advice comment
High-distress post x acknowledgment/validation
-0.8
-0.4
0.0
0.4
Coefficient on asinh(comment upvotes)

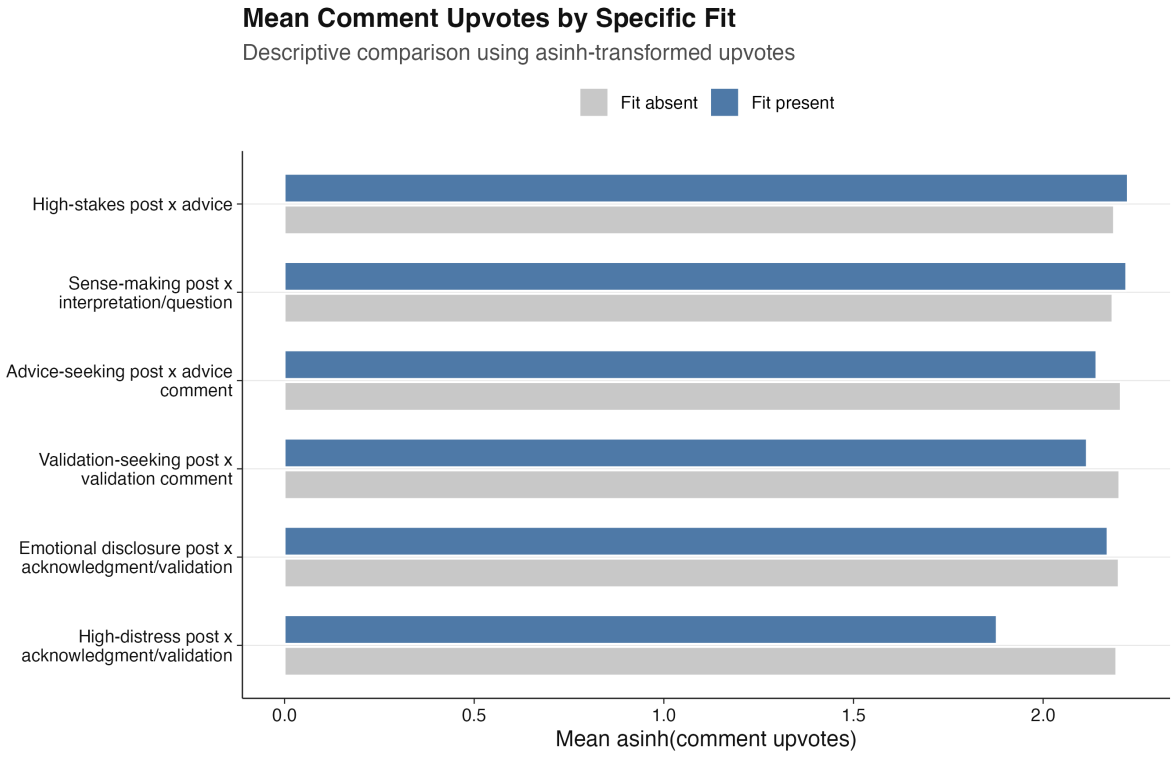

Mean Comment Upvotes by Specific Fit
Descriptive comparison using asinh-transformed upvotes
Fit absent
Fit present
High-stakes post x advice
Sense-making post x interpretation/question
Advice-seeking post x advice comment
Validation-seeking post x validation comment
Emotional disclosure post x acknowledgment/validation
High-distress post x acknowledgment/validation
0.0
0.5
1.0
1.5
2.0
Mean asinh(comment upvotes)

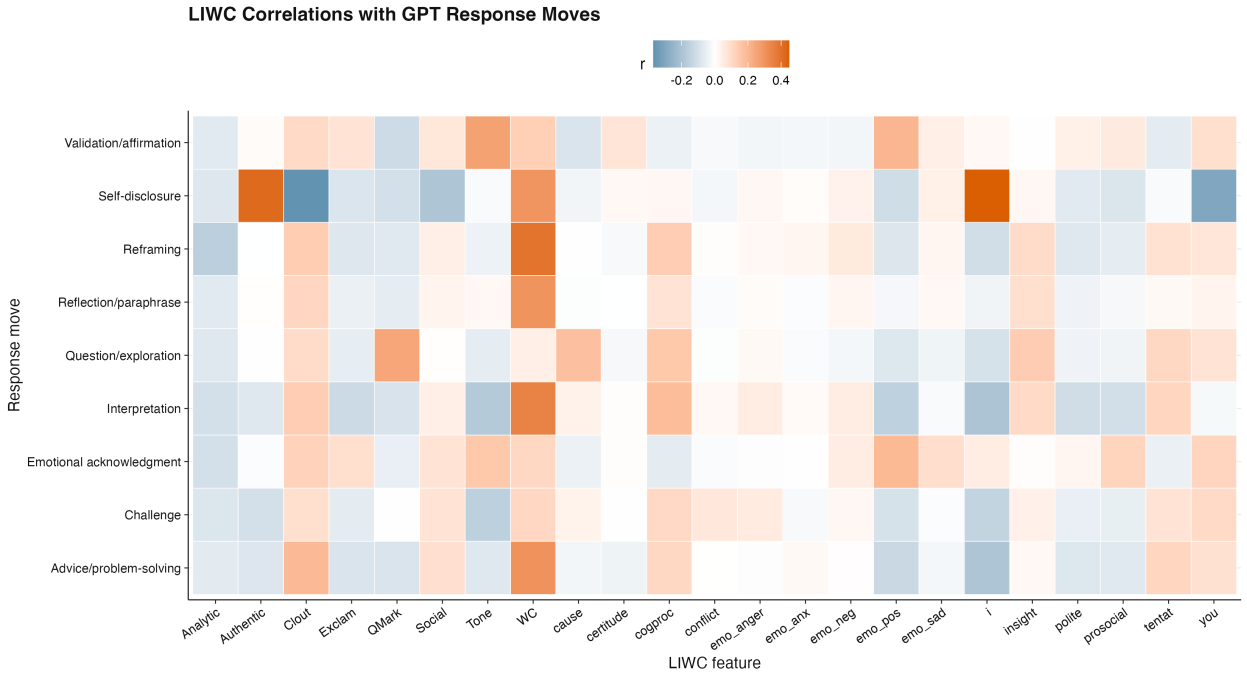
Table 12: Specific fit predicting asinh-transformed comment upvotes

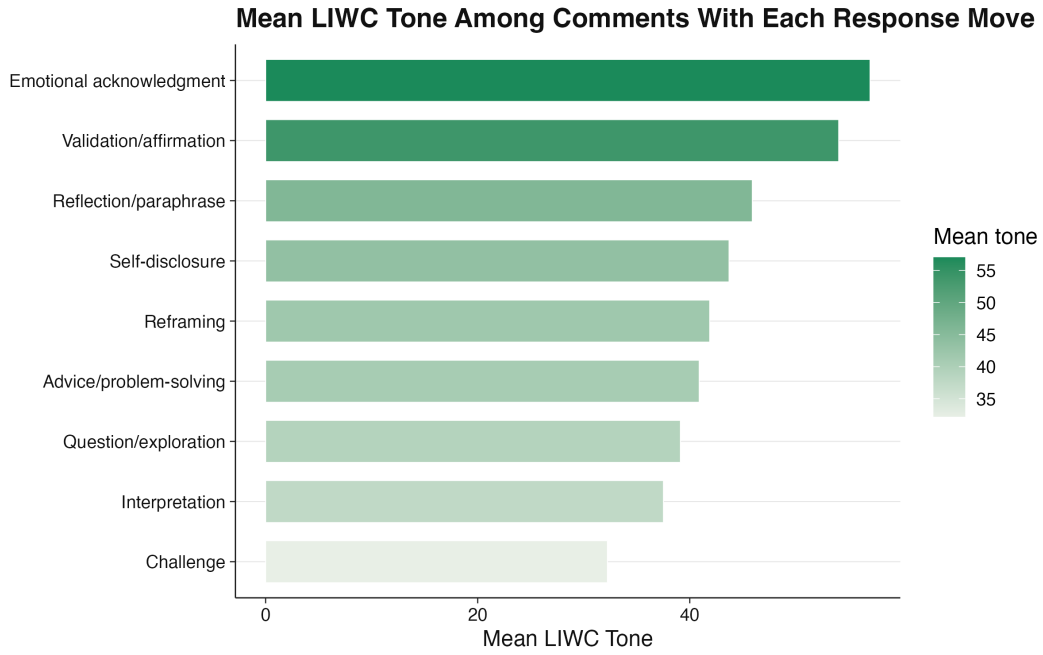
fit	estimate	95% CI	p
Advice-seeking post x advice comment	-0.105	[-0.285, 0.075]	0.254
Validation-seeking post x validation comment	-0.075	[-0.289, 0.138]	0.490
Sense-making post x interpretation/question	-0.035	[-0.229, 0.158]	0.719
Emotional disclosure post x acknowledgment/validation	0.160	[-0.002, 0.323]	0.054
High-distress post x acknowledgment/validation	-0.181	[-0.765, 0.404]	0.545
High-stakes post x advice	0.269	[-0.062, 0.599]	0.111

Overall, the upvote models provide weaker evidence than the OP uptake models. Most specific fit coefficients are small and have confidence intervals that cross zero. Emotional-disclosure fit with acknowledgment/validation is positive but borderline, and high-stakes/advice fit is positive but imprecise because those cases are relatively uncommon. Comment length and sampling strata, especially the high-engagement stratum, are stronger predictors of upvotes than the fit indicators themselves. This suggests that OP uptake may be the more sensitive outcome for testing support attunement, whereas upvotes are shaped by broader visibility, subreddit, and engagement dynamics.

LIWC and Comment Language Features

LIWC features provide a separate descriptive check on the language profile of different comment strategies. The purpose is not to validate the labels mechanically, since both LIWC and LLM labels are derived from text. Instead, LIWC helps assess whether the annotated strategies have plausible linguistic signatures. For example, advice/problem-solving comments should often contain more cognitive or action-oriented language, emotional acknowledgment should show more affective language, self-disclosure may contain more first-person language, and questions should show question-mark or interrogative patterns.





These LIWC summaries are descriptive checks on the language profile of LLM-coded response moves. They are useful for construct validation, but they should not be interpreted as independent validation of the labels because both LIWC and LLM labels are text-derived.

The most useful way to read these figures is comparatively: look for whether a response move has the expected relative profile compared with other moves, rather than treating any single LIWC coefficient as definitive. Patterns that align with expectations strengthen construct plausibility; unexpected patterns should be used to identify labels that need further human review or prompt refinement.

Qualitative Examples

The file `qualitative_examples.csv` contains selected examples for good fit with positive uptake, low fit/mismatch, question followed by elaboration, challenge/interpretation followed by pushback, and high-stakes advice.

Example 1: Good fit positive uptake

Subreddit/comment: askwomenadvice / `f1l368e`

Post: Last night my drunk boyfriend betrayed my trust during sex;As a preface, nothing remotely like this has ever happened in our relationship before, and we've been togeth…

Comment: I am so sorry. He raped you. It's not your fault. If there's anyone you trust, please go stay with them. You are not safe with this man.

OP reply: It hurts so bad to read or think that. I'm really not under a spell of some alcoholic narcissist or anything, he truly is a kind and loving person …

Example 2: Good fit positive uptake

Subreddit/comment: relationships / `jcvwyki`

Post: New friend is replacing me in my friend group.;I decided to post here because I had a panic attack this morning over this, and decided I needed to talk about it, and f…

Comment: You are not crazy for feeling this way, OP. I would feel horrible if my friends suddenly started liking my interests because of a new person. Or if they started to to hang out without me …

OP reply: Yea, maybe. I guess I just really need to think about bringing this up to them. I have a fear of coming off as manipulative and making them feel as…

Example 3: Mismatch or low fit response

Subreddit/comment: AskHR / `j7e2gha`

Post: [IL] employer won't accommodate to my life threatening allergies;My work HR won't accommodate to my perfume/cologne allergies which lead to anaphylactic shock Two of m…

Comment: This is a confusing one to me because there are very few life threatening allergies that are inhalation based. And I have 3 anaphylactic allergies if I ingest fruits or medications, but w…

Example 4: Mismatch or low fit response

Subreddit/comment: NoStupidQuestions / `ja4wdyw`

Post: I'm on a movie date to see Avatar 2 and she's totally boring or uninterested. What should I do?;She keeps checking her phone and hasn't even put her 3D Glasses on. She…

Comment: COCAINE BEAR

Example 5: Question exploration followed by elaboration

Subreddit/comment: datingoverthirty / `j2k9pp9`

Post: Profile review: 34GQ;Thanks to everyone for helping with pictures earlier! I finally got onto the apps and… I'm getting very few likes, let alone matches. Single dig…

Comment: Of the matches you are getting, do they seem like your type? Between kinky and monigamish and LGBTQ, you may just be fishing in a small pool.

OP reply: They do, so I suppose that's a good sign. LGBTQ is the biggest part of that for me, with kinky second and relationship type a more distant third. P…

Example 6: Question exploration followed by elaboration

Subreddit/comment: DecidingToBeBetter / `j3gz9z7`

Post: My girlfriend is pregnant but She's breaking up with me;So I've been with this girl let's call her Rose. She's been with me for 3 years and 6 months; a lot of things h…

Comment: Why did you not convert to islam to be able to marry her? And your parents are wrong saying if the baby is yours she would do anything to marry you. Because marrying you would mean she ha…

OP reply: It's not my parents request that I remain Christian. It's my own choice since the beginning of the relationship. I told my girlfriend and she's agr…

Example 7: Challenge or interpretation followed by pushback

Subreddit/comment: DecidingToBeBetter / `j3gz9z7`

Post: My girlfriend is pregnant but She's breaking up with me;So I've been with this girl let's call her Rose. She's been with me for 3 years and 6 months; a lot of things h…

Comment: Why did you not convert to islam to be able to marry her? And your parents are wrong saying if the baby is yours she would do anything to marry you. Because marrying you would mean she ha…

OP reply: It's not my parents request that I remain Christian. It's my own choice since the beginning of the relationship. I told my girlfriend and she's agr…

Example 8: Challenge or interpretation followed by pushback

Subreddit/comment: AskOldPeopleAdvice / `j53fdra`

Post: I am old but still need help;Hello. I am 68. I was a misery to my family as a young teenager and was self centered most of my life. I have apologized to all of them, e…

Comment: Good news ! Your almost to the end , so just be glad you are closer to the end than at the beginning! What ever makes you make it another day is great. Life is never a Carnival, its more …

OP reply: Thanks. I read plenty of books, all types.

Model Diagnostics

Table 13: Model fitting structures and diagnostics

model_type	singular	fallback_used	n
nested	FALSE	none	33
subreddit	FALSE	used subreddit	5
fixed_subreddit	NA	used fixed_subreddit	2

Most models used the preferred nested random-intercept structure. A small number used simpler fallback structures. Convergence and singularity diagnostics are saved in `model_diagnostics_convergence_log.csv`.

References

Alghamdi, Z., Kumarage, T., Agrawal, G., Karami, M., Almuteb, I., & Liu, H. (2025). *RedditESS: A Mental Health Social Support Interaction Dataset – Understanding Effective Social Support to Refine AI-Driven Support Tools*. arXiv. https://arxiv.org/abs/2503.21888

Barbee, A. P., & Cunningham, M. R. (1995). An experimental approach to social support communications: Interactive coping in close relationships. *Annals of the International Communication Association*.

Cutrona, C. E., & Suhr, J. A. (1992). Controllability of stressful events and satisfaction with spouse support behaviors. *Communication Research*.

Goldsmith, D. J. (2004). *Communicating Social Support*. Cambridge University Press.

Gueorguieva, E., Zhan, H., Suh, J., Hernandez, J., Lau, T., Li, J. J., & Ong, D. C. (2026). *AI generates well-liked but templatic empathic responses*. arXiv. https://arxiv.org/abs/2604.08479

Caveats and Next Steps

1. These analyses use LLM-coded annotations, not final human-validated labels.
2. The completed sample is 2765 rows out of the planned 3,000 because the API run stopped for insufficient quota.
3. The sample is enriched, so random-subset results should be used when population-like descriptive checks are needed.
4. Main/stable labels should be prioritized in reporting: advice-seeking, emotional disclosure, validation/appraisal, sense-making; comment validation, interpretation, question/exploration, advice/problem-solving, self-disclosure; OP gratitude and elaboration.
5. Exploratory/caution labels include high distress, high stakes, perspective-sharing boundary flag, reflection/paraphrase, reframing, challenge, and low-frequency OP uptake labels.

Recommended next step: inspect the qualitative examples and model diagnostics, then decide whether to finish the remaining 235 annotations or proceed with the completed 2,765-row preliminary sample for internal reporting.